



what's spoppin'?

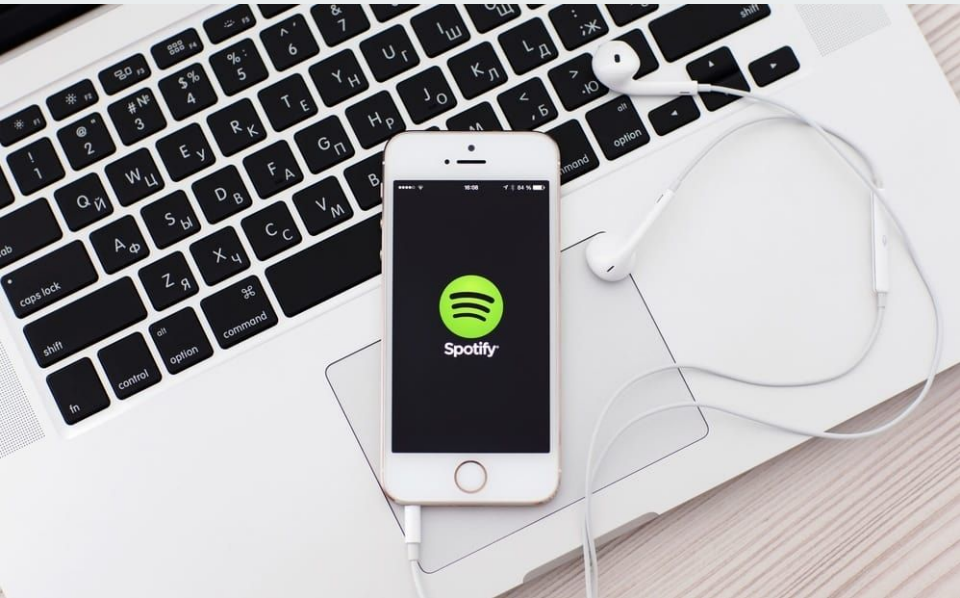
Analysis & Prediction of Popularity of Spotify Tracks
By Erin Vu





Overview

1. Business Problem
2. Data Overview
3. Modeling Process
4. Model Results
5. Conclusion



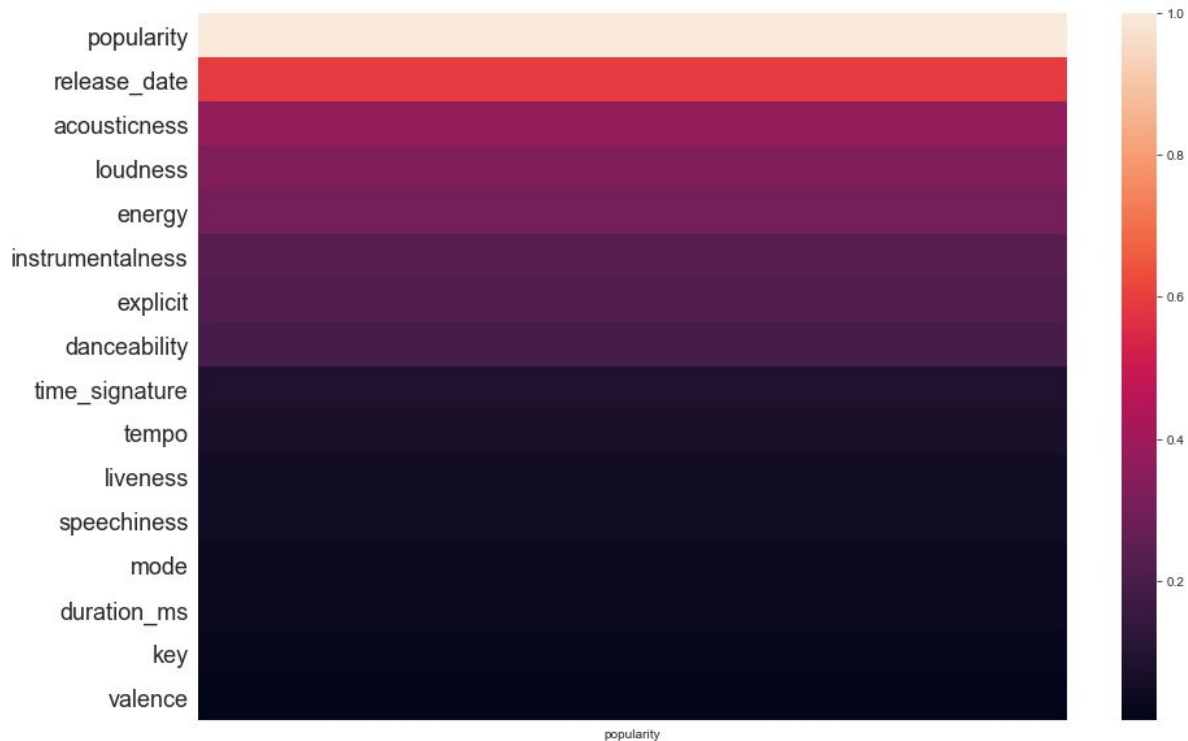
Business Problem

- Stakeholder: Music Production Company
- Dominate streaming/radio charts
- Creating popular songs
- Understanding popular song features
- Using features to predict popularity

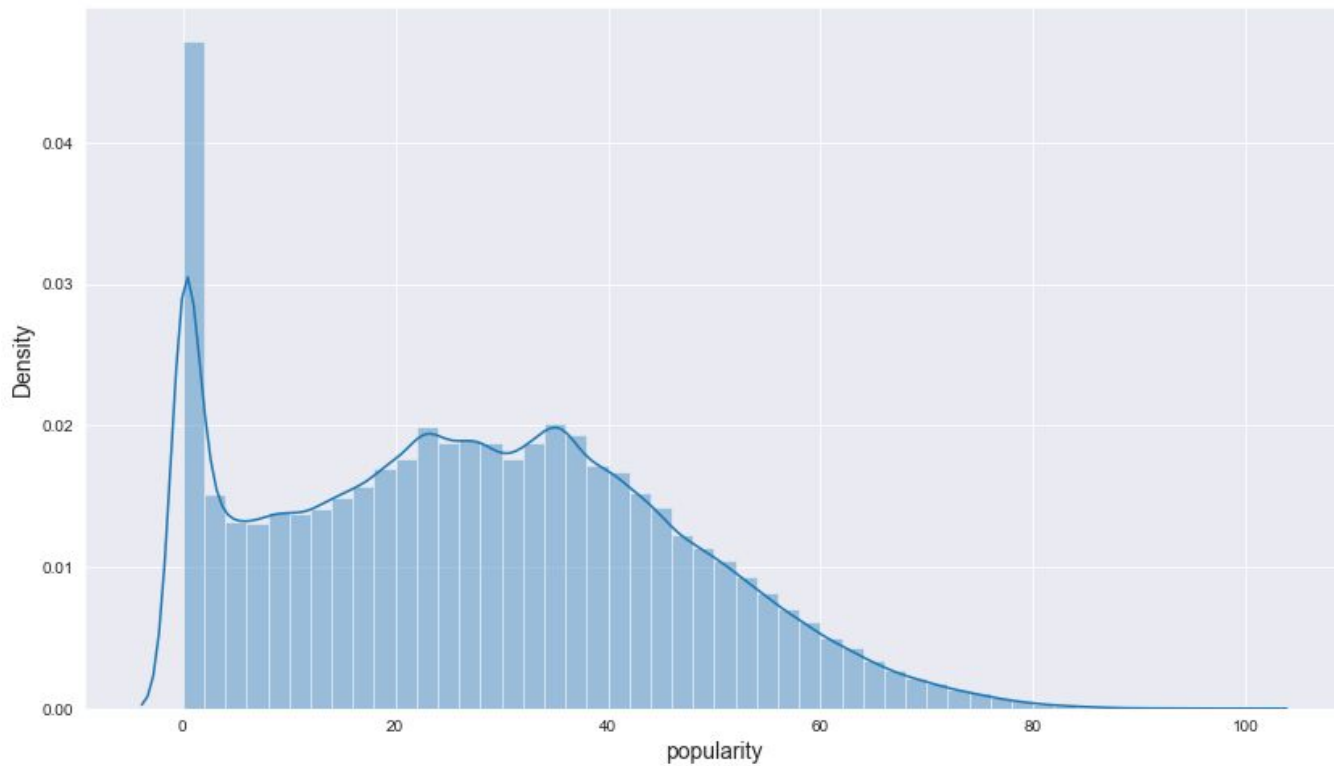


Data

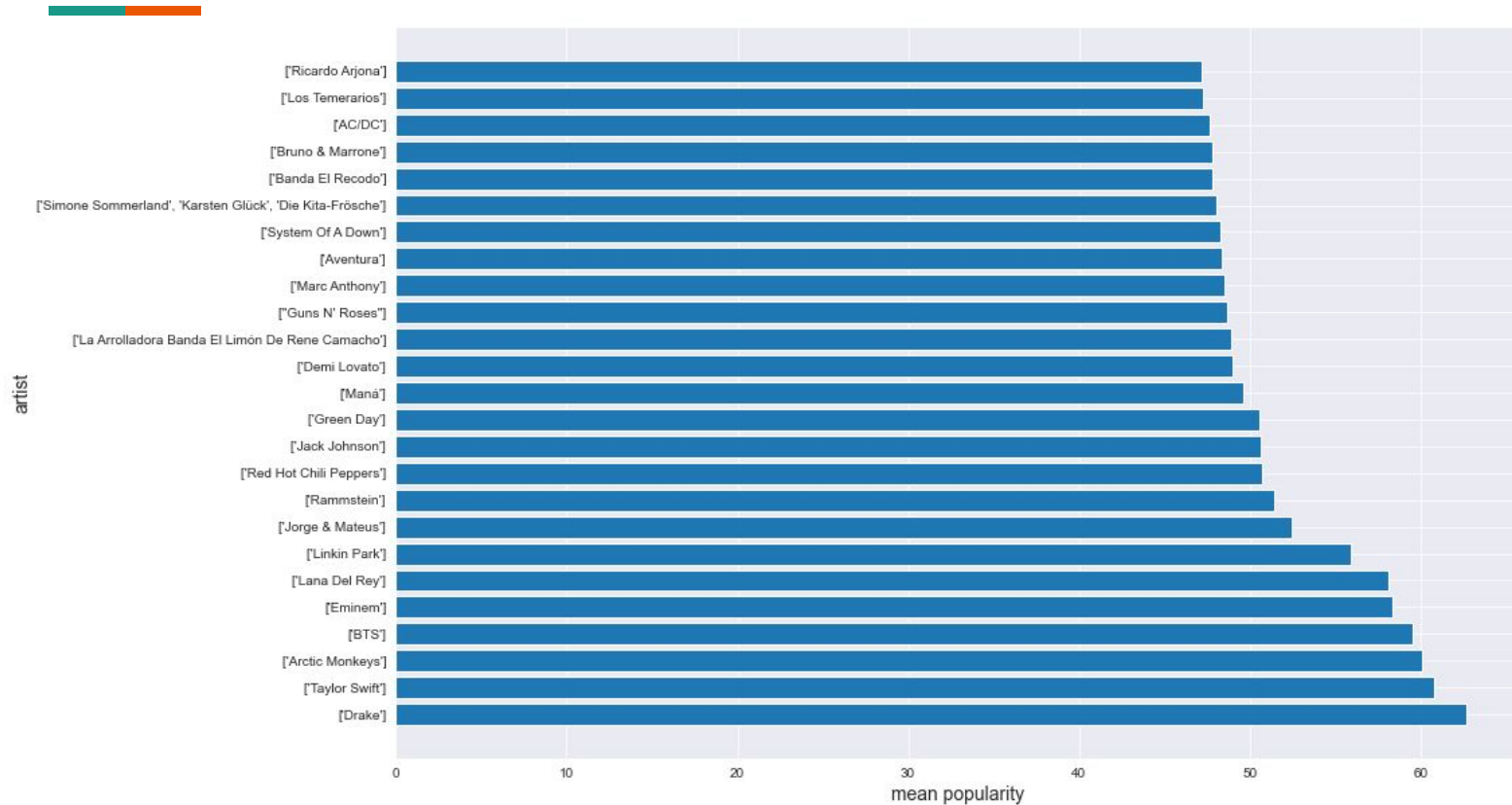
- Spotify Kaggle Dataset
- 1922-2021
- 600k songs
- 20 Features

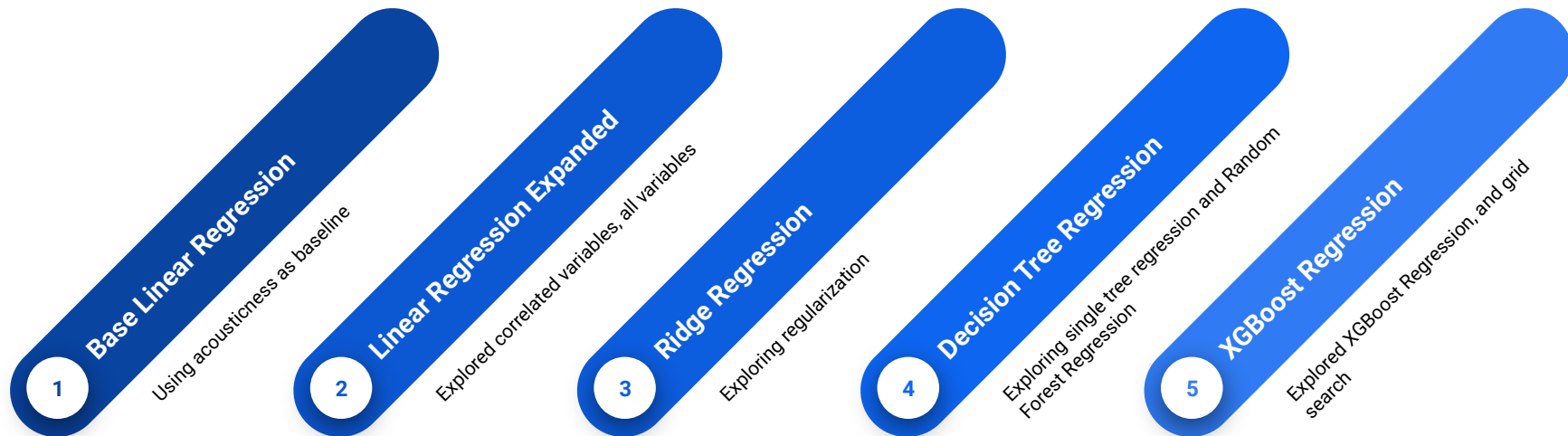


Target: Popularity



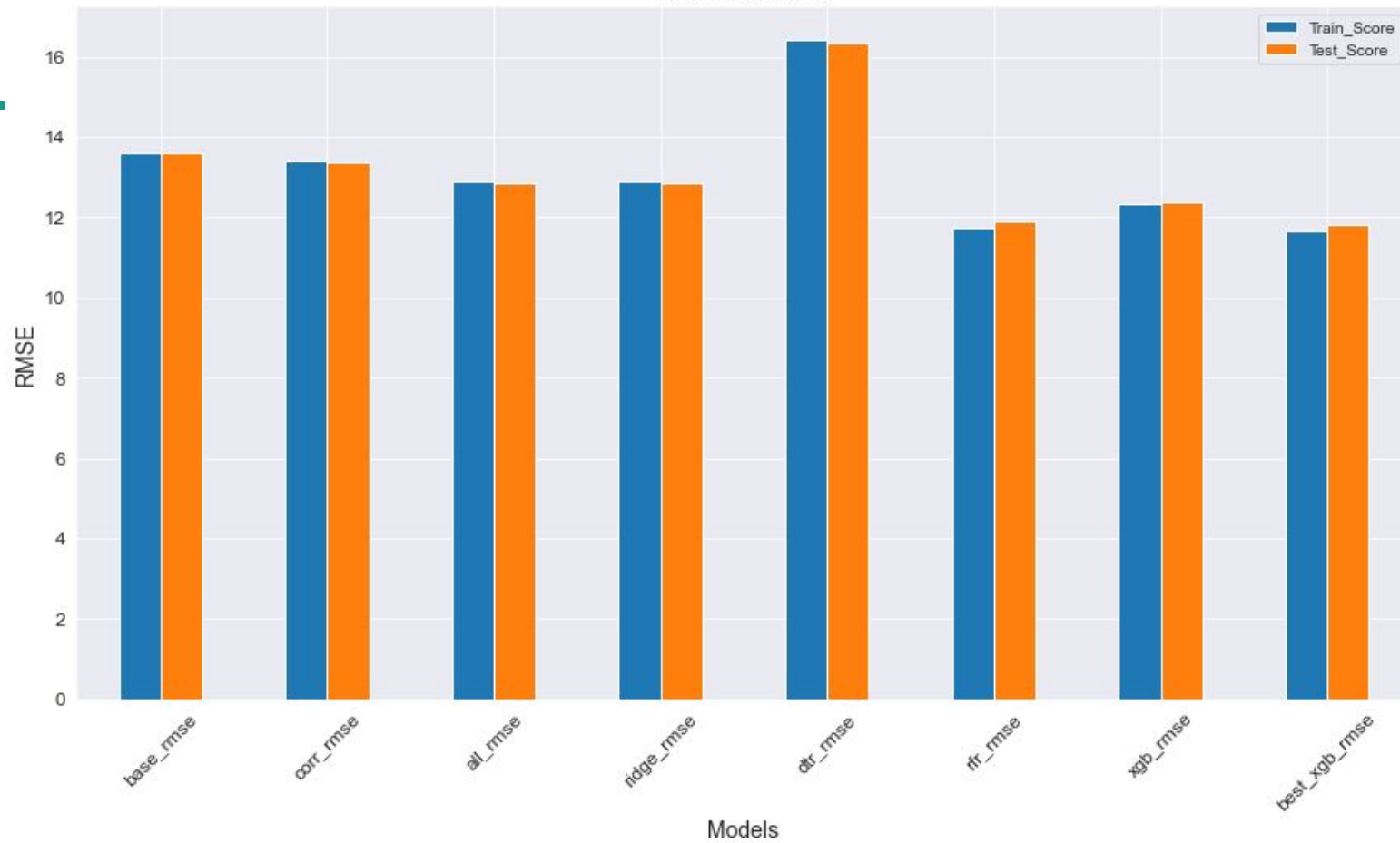
Top 25 Artists with > 100 Songs



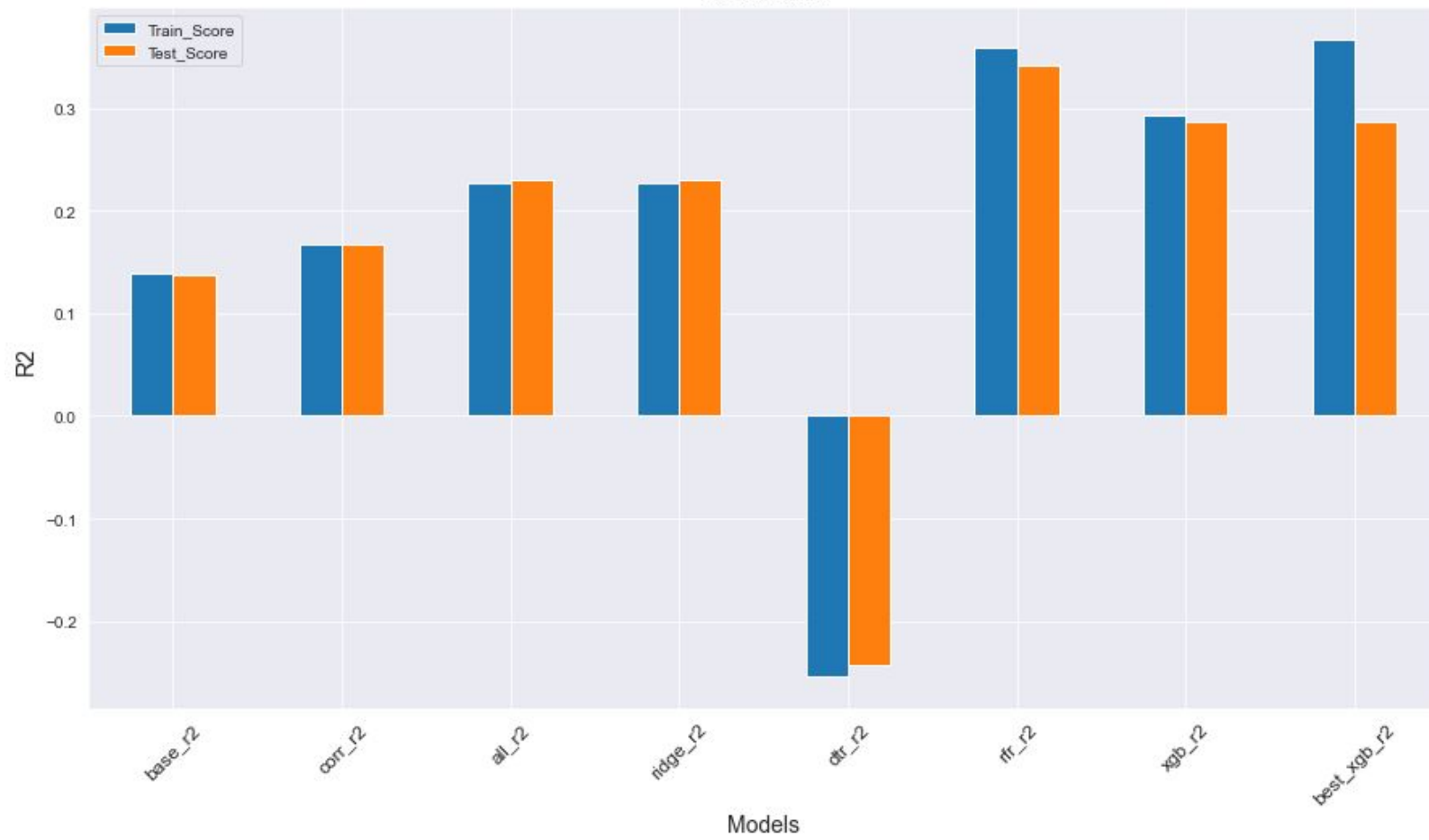


Modeling Process

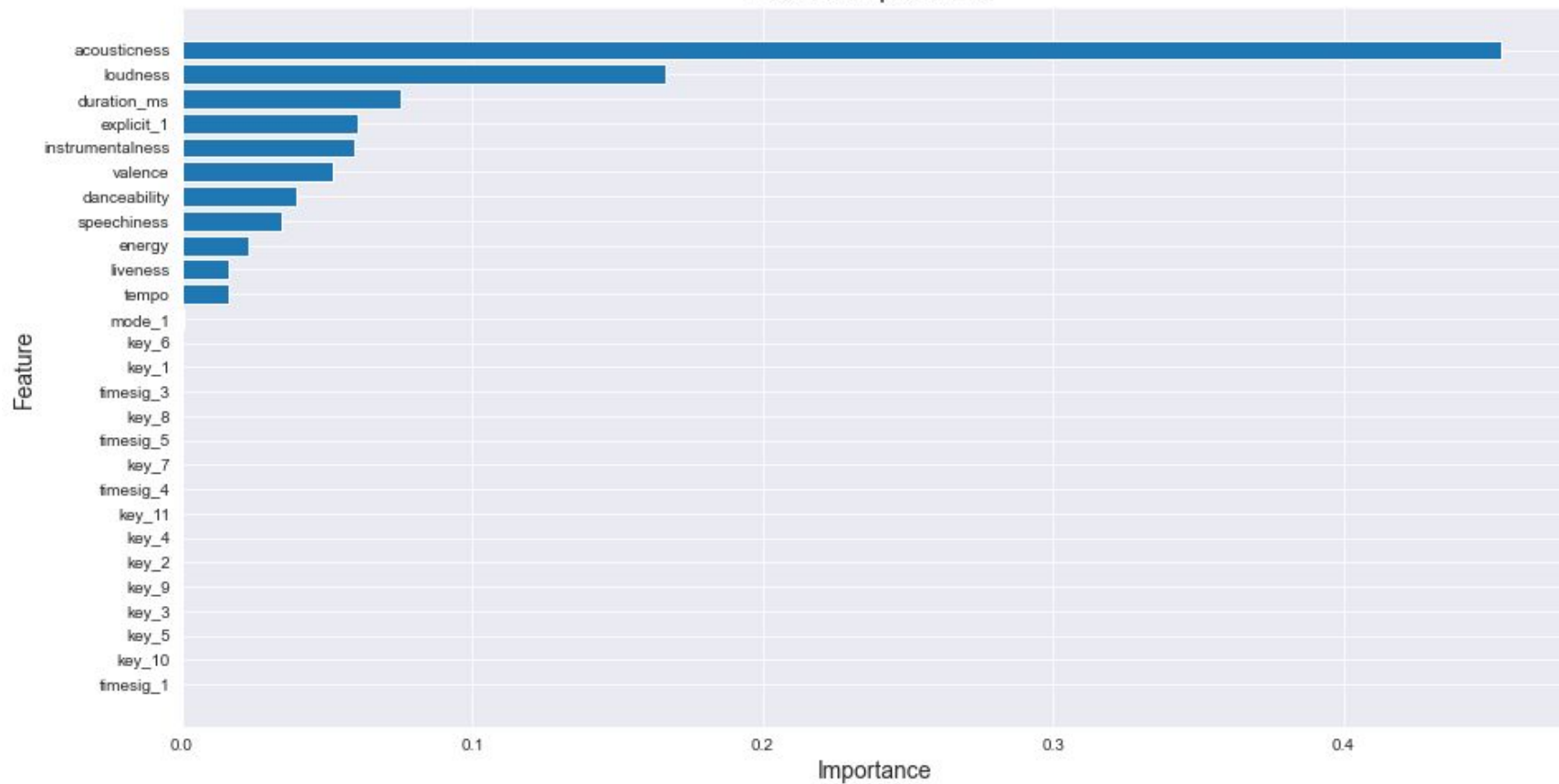
RMSE Scores



R2 Scores



Feature Importances





Conclusion and Next Steps

- Best Model: Random Forest Regression
- Stray away from acoustic/low volume releases
- Different types of music
 - Model to predict within genres
 - Subset different time periods

Thank You

Erin Vu

<https://github.com/ekvu>

Erin.vu94@gmail.com

